

# A meaningful expansion around detailed balance

Matteo Colangeli,<sup>1</sup> Christian Maes,<sup>2</sup> and Bram Wynants<sup>3</sup>

<sup>1</sup>*Dipartimento di Matematica, Politecnico di Torino,  
Corso Duca degli Abruzzi 24, 10129 Torino, Italy\**

<sup>2</sup>*Instituut voor Theoretische Fysica, K. U. Leuven, 3001 Leuven, Belgium*

<sup>3</sup>*Institut de Physique Théorique, CEA-Saclay,  
F-91191 Gif-sur-Yvette Cedex, France*

## Abstract

We consider Markovian dynamics modeling open mesoscopic systems which are driven away from detailed balance by a nonconservative force. A systematic expansion is obtained of the stationary distribution around an equilibrium reference, in orders of the nonequilibrium forcing. The first order around equilibrium has been known since the work of McLennan (1959), and involves the transient irreversible entropy flux. The expansion generalizes the McLennan formula to higher orders, complementing the entropy flux with the dynamical activity. The latter is more kinetic than thermodynamic and is a possible realization of Landauer's insight (1975) that, for nonequilibrium, the relative occupation of states also depends on the noise along possible escape routes. In that way nonlinear response around equilibrium can be meaningfully discussed in terms of two main quantities only, the entropy flux and the dynamical activity. The expansion makes mathematical sense as shown in the simplest cases from exponential ergodicity.

PACS numbers: 05.40.-a, 05.70.Ln, 05.10.Gg

---

\*Electronic address: colangeli@calvino.polito.it

## I. INTRODUCTION

Recent years have seen an intensive search for a theoretical framework underlying nonequilibrium fluctuations. The nature of nonequilibrium is diverse and rich but we hope to uncover some unifying structures that, ideally, would give an extension of the Gibbs formalism. Still, today, the mathematical models are very often simplified to Markov dynamics for jump or diffusion processes that represent open systems in weak contact with one or more reservoirs. These systems (such as molecular configurations) can be very small but nevertheless we have learnt how to attach thermodynamic meaning to various path-dependent quantities, see e.g. [27, 33], as they relate to what happens in the large environment. Most of that has concentrated around the concepts of energy and entropy, so that much of standard irreversible thermodynamics also got a formulation for mesoscopic systems. Another question has concerned response theory for these Markov dynamics and some systematics have been obtained there also, e.g. in [1, 2, 22, 29, 31, 35]. A related issue concerns the characterization of the stationary distribution for a dynamics that breaks the condition of detailed balance. Of course we need a physically meaningful violation of detailed balance, and a useful interpretation is given by the condition of local (sometimes called, generalized) detailed balance, [5, 11, 18, 19, 34]. The latter requires that the power dissipated to the environment during a transition equals the logarithmic ratio between forward and backward transition rates. From there one can hope to find the corresponding stationary distribution in terms of these irreversible entropy fluxes. That is exactly what was achieved by McLennan in 1959, [30]. In that way, the stationary distribution picks up thermodynamic information and is no longer “just a stationary solution” of the Master equation. However, that McLennan proposal only works close-to-equilibrium [12–14, 26]. There does not appear to be a ready extension beyond the linear regime in terms of entropy considerations only. The present paper takes a next step but we need to go beyond the purely entropic concepts we are used to from heterogeneous equilibrium. That is, to go to second and higher order in an expansion around detailed balance we need another concept to complement the entropy fluxes. That novel quantity is called the dynamical activity and is much related to the notion of escape rate: it measures the reactivity and instability of a trajectory. Dynamical activity is thus much more concerned with kinetics than it is embedded into thermodynamics but by introducing it, we can complete the expansion beyond linear order in the nonconservative forces around equilibrium. In that sense, we add to the spirit of the McLennan proposal the insights of Landauer and others that for nonequilibria, the noise behavior along in-

and outgoing trajectories enters critically into the determination of state plausibilities, [15–17].

In the expansion of the stationary distribution, every term at any order in the nonequilibrium forcing, just contains the same (dynamical) observables in various combinations of time-correlation functions under the reference equilibrium process. For example, in some precise sense correct up to third order we get

$$\rho(x) = \rho_o(x) \left[ 1 - \langle S \rangle_x^o + \frac{1}{2} \langle S \mathcal{T}_1 \rangle_x^o - \frac{1}{24} \langle S (S^2 + 3\mathcal{T}_1^2 - 12\mathcal{T}_2) \rangle_x^o \right]$$

for the stationary distribution  $\rho$  on states  $x$  in terms of the equilibrium distribution  $\rho_o$ . The averages  $\langle \cdot \rangle_x^o$  are over the detailed balance process started from state  $x$  while  $S$  is the irreversible entropy flux and  $\mathcal{T}_1, \mathcal{T}_2$  denote the first and second order to the dynamical activity. These path observables  $S$  and  $\mathcal{T}_{1,2}$  depend of course on the nonequilibrium dynamics:  $S$  is basically determined by the work done by the nonconservative force in a particular trajectory, and  $\mathcal{T}_{1,2}$  measures the expected dynamical activity long the followed trajectory. Specific details, rewriting and mathematical precision follow below.

In the next section, we specify the Markovian set-up and we define the various objects such as  $\rho, S$  and  $\mathcal{T}$  in the above. We concentrate on overdamped diffusions and jump processes for introducing (in a particular way) the nonequilibrium driving. Sections III–IV–V give the main idea and the structure of the expansion with some writing out for specific models to lowest order. The last section about nonlinear response suggests some immediate application before we conclude the paper. Appendix B defines the entropy flux and the dynamical activity for underdamped diffusions.

## II. SET-UP OF MARKOV STOCHASTIC DYNAMICS

We restrict our analysis to Markovian stochastic models for mesoscopic systems driven by a nonconservative force, i.e., a force that cannot be derived from a potential. We imagine such systems to be immersed in an environment in thermal equilibrium at some inverse temperature  $\beta$ . All conditions are time-independent and we want to characterize the statistical distribution of states for our open system when reaching stationarity.

A first standard choice is to consider overdamped diffusions for state  $x \in \mathbb{R}^n$ , according to which

$$\dot{x}_t = \chi \cdot F(x_t) + \sqrt{2\frac{\chi}{\beta}} \xi_t, \quad \xi_t = \text{standard white noise} \quad (1)$$

where the mobility  $\chi$  is a positive definite  $n \times n$ -matrix not depending on  $x$  (for simplicity only). The total force equals

$$F = \epsilon f - \nabla U$$

for the energy  $U$  of the system, and with  $f$  the nonconservative force with amplitude  $\epsilon$  (our small number). When  $\epsilon = 0$ , the dynamics satisfies the condition of detailed balance, and given enough time the distribution of states converges to the equilibrium distribution  $\rho_o(x) \propto e^{-\beta U(x)}$ , so we assume.

Whenever  $\epsilon \neq 0$  the system is not in equilibrium. The Fokker-Planck equation for the evolution of distributions  $\mu_t$  is

$$\frac{\partial \mu_t}{\partial t}(x) + \nabla \cdot [\chi F(x) \mu_t(x) - \frac{\chi}{\beta} \nabla \mu_t(x)] = 0$$

The stationary solution  $\rho$  thus satisfies  $\beta \nabla \cdot F(x) \rho(x) = \Delta \rho(x)$ , but we would be much helped by further more explicit, physical or systematic understanding of that  $\rho$ . One option is to find an expansion for that stationary distribution  $\rho(x)$  of the system in orders of  $\epsilon$ , assuming uniformly and exponentially fast relaxation behavior. That is the programme of the present paper and we will find that the main quantities in such an expansion are directly related to two specific path observables, that we now introduce.

We fix a large time-interval  $[0, T]$  and we consider paths  $\omega = (x_t, t \in [0, T])$ , on which we define

$$\begin{aligned} \mathcal{T}(\omega) &= \frac{\epsilon^2 \beta}{2} \int_0^T dt f \cdot \chi f - \epsilon \beta \int_0^T dt f \cdot \chi \nabla U + \epsilon \int_0^T dt \chi \nabla \cdot f \\ S(\omega) &= \epsilon \beta \int_0^T dx_t \circ f(x_t) \end{aligned} \quad (2)$$

The last stochastic integral (with the  $\circ$ ) is in the sense of Stratonovich; in that way  $S$  is identified with  $\beta$  times the work done by the nonconservative force — for short, we speak about the entropy flux  $S$ . The quantity  $\mathcal{T}$  is less familiar, and it contains both order  $\epsilon$  and order  $\epsilon^2$ . Correspondingly, we write  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2$  for the first and second order. Its meaning is best understood before the continuum limit, in terms of jump processes to which we turn next.

We consider a Markov jump process with discrete states  $x$  and as a reference we take jump rates  $k_o(x, y)$  for the transition  $x \rightarrow y$  which satisfy detailed balance, i.e.,

$$k_o(x, y) e^{-\beta U(x)} = e^{-\beta U(y)} k_o(y, x)$$

In this case the stationary Master equation is solved by  $\rho_o(x) \propto \exp(-\beta U(x))$ . To this reference we add an extra flux  $f(x, y) = -f(y, x)$  of energy in the transition  $x \rightarrow y$ , and write

$$k_\epsilon(x, y) = k_o(x, y) e^{\frac{\beta \epsilon}{2} f(x, y)} \quad (3)$$

To establish nonequilibrium we ask that the fluxes  $f(x, y)$  cannot all be rewritten as the difference  $V(x) - V(y)$  of a unique potential  $V$ , which means that there are loops  $x_1 \rightarrow x_2 \rightarrow \dots x_n = x_1$  over which the sum  $f(x_1, x_2) + f(x_2, x_3) + \dots f(x_{n-1}, x_n) \neq 0$  does not vanish. In terms of the condition of local detailed balance, these fluxes  $f(x, y)$  should be interpreted as the product of a displacement of a certain quantity and a nonconservative force, see [11], but here we do not need this formulation. The  $\epsilon$  is the magnitude of the nonequilibrium forcing. Finally, in (3) we have chosen to omit an extra symmetric prefactor  $\psi_\epsilon(x, y)$  because the more general choice for (3) would be  $k_\epsilon(x, y)/k_o(x, y) = \psi_\epsilon(x, y) \exp[\beta\epsilon f(x, y)/2]$ . We take however  $\psi_\epsilon(x, y) = \psi_\epsilon(y, x) = 1$  for much greater simplicity.

Again, for  $\epsilon \neq 0$ , the stationary probability law  $\rho$  is only known indirectly as solution of the Master equation  $\sum_y [\rho(y)k(y, x) - \rho(x)k(x, y)] = 0$  for all  $x$ .

This time the entropy flux is

$$S(\omega) = \epsilon\beta \sum_{t \leq T} f(x_{t-}, x_t) \quad (4)$$

as a sum over the jump times  $t$  in the trajectory  $\omega = (x_s, s \in [0, T])$ , and the dynamical activity is

$$\mathcal{T}(\omega) = 2 \int_0^T dt \sum_y k_o(x_t, y) [e^{\frac{\beta\epsilon}{2} f(x_t, y)} - 1] \quad (5)$$

Now we see better where the name activity comes from:  $\mathcal{T}(\omega)$  is the difference in the escape rates, integrated over the trajectory. The escape rate  $\sum_y k(x, y)$  measures the frequency by which the system exits state  $x$ , and in that way it counts the expected number of transitions away from  $x$ . In other words,  $\mathcal{T}(\omega)$  sees how the escape rate away from the trajectory  $\omega$  changes when adding the forcing  $f$ . In the appropriate rescaling the expression (2) is simply the continuum limit of (5) from Markov jump to (overdamped) diffusions. We do that computation in Appendix A; the case of underdamped or inertial diffusions is shortly discussed in B. We repeat that we use here (3), i.e., with prefactor  $\psi_\epsilon = 1$ ; otherwise the expression for the activity gets more complicated — the major part of the analysis would however remain unchanged.

### III. EXPANSION: MAIN IDEA

The expansion of the stationary distribution starts from a simple idea which was applied already in [7, 12, 20, 26]: the single-time distributions on states, and in particular the stationary distribution, can be obtained from its embedding in the path space distribution. The latter is the

distribution on the level of trajectories or paths  $\omega$  and gives the weight  $P(\omega)$  for path-integrals.  $P$  is much more directly obtained and is much better-behaved than its projections on single time layers. In fact, we can give explicit expressions for the “action”  $A(\omega)$  in

$$P(\omega) = e^{-A(\omega)} P^o(\omega) \quad (6)$$

that connects the distribution  $P$  on paths starting from  $\rho_o$  but *with* driving  $f$ , with the full equilibrium reference distribution  $P^o$ , see e.g. [28] for some useful techniques. The action  $A$  is typically local in space-time and thus is similar to Hamiltonians or Lagrangians that we meet in (equilibrium statistical) mechanics, see e.g. [23, 27]. We can verify that

$$A = (\mathcal{T} - S)/2$$

as defined above for overdamped and jump processes. Furthermore, the action  $A$  in (6) is left unchanged when both processes start from the same state  $x$ ,

$$P_x(\omega) = e^{-A(\omega)} P_x^o(\omega) \quad (7)$$

Now comes the embedding. As defined before, both nonequilibrium and equilibrium processes in (6) start at time  $t = 0$  from data distributed with the equilibrium  $\rho_o$ . At time  $T$  later, the probability to find the driven system in state  $x$  is

$$p(x, T) = \langle \delta(x_T - x) \rangle_{\rho_o}^\epsilon$$

averaging over the trajectories for the nonequilibrium dynamics. We assume that for  $T \uparrow +\infty$  there is exponentially fast convergence  $p(x, T) \rightarrow \rho(x)$  (in the sense of densities) to the stationary distribution of the nonequilibrium process, uniformly in  $\epsilon$ . By (6) we can rewrite

$$p(x, T) = \langle \delta(x_T - x) e^{-A(\omega)} \rangle_{\rho_o}^o \quad (8)$$

which is now an expectation value for the full equilibrium (detailed balance) process (and then we omit the subscript  $\rho_o$ ). By time-reversal invariance, (8) equals

$$p(x, T) = \langle \delta(x_0 - \pi x) e^{-A(\theta\omega)} \rangle_{\rho_o}^o = \rho_o(x) \langle e^{-A(\theta\omega)} \rangle_{x_0=\pi x}^o \quad (9)$$

where the time-reversal operator  $\theta$  acts as

$$\theta\omega = ((\pi x)_{T-t}, 0 \leq t \leq T)$$

with  $\pi x$  equal to  $x$  except for flipping the velocities (if they are part of the state-description) or other variables with negative parity under time-reversal. In equilibrium  $\rho_o(\pi x) = \rho_o(x)$ . Equality (9) can still be rewritten in terms of  $S$  and  $\mathcal{T}$  getting

$$p(x, T)\rho_o(x) \left\langle e^{-(S+\mathcal{T})/2} \right\rangle_{x_0=\pi x}^o \quad (10)$$

Indeed, the decomposition  $A = (\mathcal{T} - S)/2$  follows the symmetry under time-reversal,

$$S(\omega) = A(\theta\omega) - A(\omega)$$

$$\mathcal{T}(\omega) = A(\theta\omega) + A(\omega)$$

The fact that the quantity  $A(\theta\omega) - A(\omega)$  is the excess entropy flux from the system into the environment during the process  $\omega$  is one of the main discoveries for the construction of nonequilibrium statistical mechanics of the last decade, see [4, 25] and the references in e.g. [8, 23, 29]. Excess means the difference between the nonequilibrium process and the reference equilibrium process. Specifically, this excess is here equal to the work done by the nonequilibrium force  $\epsilon f$ , multiplied by  $\beta$ . The dynamical activity  $\mathcal{T}$  has been introduced and used before [21, 24] but has no thermodynamic tradition. Its role is kinetic and that it influences the relative stability of states was somehow emphasized long before, cf. [15–17].

The left-hand side of (10) is assumed to converge exponentially fast to the stationary law  $\rho(x)$ , but there is a problem with its right-hand side because both  $S$  and  $\mathcal{T}$  are time-extensive of order  $T$ , being sent to infinity. We will therefore need to control the limit  $T \uparrow +\infty$  and to worry about the exchange with the sum of the perturbation series. An important point here that follows from (7), is the normalization

$$\left\langle e^{(S-\mathcal{T})/2} \right\rangle_x^o = 1 \quad (11)$$

valid under the equilibrium process but started from an arbitrary state  $x$ . That itself can be expanded in orders of  $\epsilon$ , and takes care of many cancellations. Finally, one must use that  $S$  is anti-symmetric, and  $\mathcal{T}$  is symmetric under time-reversal, so that over all time-intervals  $[0, T]$

$$\langle S^n \mathcal{T}^m \rangle^o = 0, \quad n \text{ odd} \quad (12)$$

for all  $m$  and for all odd powers  $n$ . With these ingredients (10)–(11)–(12), combined with fast relaxation for the equilibrium process, all is in place to start a systematic expansion. In the particular cases we have in mind, see Section II, the path function  $S$  is simply first order in  $\epsilon$  and  $\mathcal{T}$  is

either  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2$  second order in  $\epsilon$  for diffusions, see (2), or of arbitrary order  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 + \dots$  for Markov jump processes. For practical matters our expansion including second or third order around equilibrium is already new and relevant.

#### IV. GENERAL EXPANSION

Formally, expanding (10) just gives

$$\frac{p(x, T)}{\rho_o(x)} = 1 - \frac{1}{2} \langle S + \mathcal{T}_1 \rangle_{x_0=\pi x}^o + \frac{1}{2} \left\langle -\mathcal{T}_2 + \frac{(S + \mathcal{T}_1)^2}{4} \right\rangle_{x_0=\pi x}^o + O(\epsilon^3)$$

On the other hand, (11) gives

$$0 = \frac{1}{2} \langle S - \mathcal{T}_1 \rangle_{x_0=\pi x}^o, \quad 0 = \left\langle \frac{(S - \mathcal{T}_1)^2}{8} - \frac{\mathcal{T}_2}{2} \right\rangle_{x_0=\pi x}^o, \dots$$

Adding or subtracting these relations from the corresponding orders of the expansion of the stationary distribution simplifies matters. In the end, in every order of the expansion we can choose that only those terms survive which are averages of quantities antisymmetric in time,

$$\begin{aligned} \frac{p(x, T)}{\rho_o(x)} = & 1 - \langle S \rangle_{x_0=\pi x}^o + \frac{1}{2} \langle S \mathcal{T}_1 \rangle_{x_0=\pi x}^o \\ & - \frac{1}{24} \langle S(S^2 + 3\mathcal{T}_1^2 - 12\mathcal{T}_2) \rangle_{x_0=\pi x}^o + O(\epsilon^4) \end{aligned} \quad (13)$$

The above considerations can be systematized. The formal expansion that results, after also taking into account the normalization in (11), is

$$p(x, T) = \rho_o(x) \left[ 1 + \sum_{m=1}^{\infty} \left\langle B_m \left( -\frac{S + \mathcal{T}}{2} \right) - B_m \left( \frac{S - \mathcal{T}}{2} \right) \right\rangle_{x_0=\pi x}^o \right] \quad (14)$$

where we introduced a shorthand notation for the path-dependent functionals  $B_m(G)$  acting on path observables  $G(\omega, \epsilon)$ , defined via

$$\begin{aligned} B_m(G) &= B_m(G_1, \dots, G_{m-k+1}) \\ &= \sum_{k=1}^m \sum_{\sigma} \frac{1}{b_1! \dots b_{m-k+1}!} \left( \frac{G_1}{1!} \right)^{b_1} \dots \left( \frac{G_{m-k+1}}{(m-k+1)!} \right)^{b_{m-k+1}} \end{aligned} \quad (15)$$

for  $G_\ell(\omega) = d^\ell G / d\epsilon^\ell(\omega, 0)$ . The sum in (15) extends over all sequences  $\sigma$  of non-negative coefficients  $\sigma = (b_1, \dots, b_{m-k+1})$ , such that  $\sum_{j=1}^{m-k+1} b_j = k$  and  $\sum_{j=1}^{m-k+1} j b_j = m$ . Note that the  $B_m$  are versions of the so called complete Bell polynomials. There exist alternative, although

equivalent, expressions for (14), see e.g. [9]. On the other hand, the expression we use has the advantage of being compact and suitable for numerical implementation.

From (14) and (15), we can write the explicit expression for the  $m$ -th order in  $\epsilon$ . Remember that we write  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 + \dots$  and that  $S$  is of order  $\epsilon$  while  $\mathcal{T}_n$  is of order  $\epsilon^n$ . The result is

$$\frac{p(x, T)}{\rho_o(x)} = 1 + 2 \sum_{m=1}^{\infty} \sum_{\sigma_m} \left(\frac{-1}{2}\right)^{\sum b_j} \frac{1}{b_0! b_1! \dots b_m!} \left\langle S^{b_0} \mathcal{T}_1^{b_1} \mathcal{T}_2^{b_2} \dots \mathcal{T}_m^{b_m} \right\rangle_{x_0=\pi x}^o \quad (16)$$

The sum in (16) extends over all sequences  $\sigma_m$  of non-negative integers  $(b_0, b_1, \dots, b_m)$ , such that  $b_0$  is odd and  $b_0 + \sum_{j=1}^m j b_j = m$ . For illustration to construct (13),  $m = 1$  requires  $b_0 = 1$  and all other  $b_j = 0$ ;  $m = 2$  requires  $b_0 = b_1 = 1$  with all the other  $b_j = 0$ ;  $m = 3$  allows three cases  $b_0 = b_2 = 1, b_0 = 1, b_1 = 2$  and  $b_0 = 3$  each time with all other  $b_j = 0$ .

In the case of diffusions (where  $p(x, T)$  must be understood as a probability density with respect to  $dx$ ), see (2), we have  $\mathcal{T}_n = 0$  for  $n > 2$  so that we must then also require  $b_j = 0, j \geq 3$  in each  $\sigma_m$ .

The first important thing to observe about the expansion (16) is that it converges for fixed state  $x$  and uniformly in time  $T \uparrow \infty$  if there is  $c = c(x) < \infty$  such that each term is bounded like

$$\left| \left\langle S^{b_0} \mathcal{T}_1^{b_1} \mathcal{T}_2^{b_2} \dots \mathcal{T}_m^{b_m} \right\rangle_{x_0=\pi x}^o \right| \leq \epsilon^m c^{\sum b_j} \quad (17)$$

The reason is that

$$\sum_{m=1}^{\infty} \epsilon^m \sum_{\sigma_m} \left(\frac{c}{2}\right)^{\sum b_j} \frac{1}{b_0! b_1! \dots b_m!} \leq \sum_{m=1}^{\infty} \epsilon^m \sum_{k=1}^m \frac{c^k (m+1)^k}{2^k k!}$$

converges for small enough  $\epsilon$ .

We only argue for (17) explicitly for the first and the second order. The first order is the McLennan formula (see immediately below) and has been treated before in [26]. The second order adds a new complication (to be treated below) and that complication is repeated for the higher order terms and can be solved in the same way. At any rate we cannot quite leave it with (13) or with (16) because we are interested in the limit  $T \uparrow \infty$  and  $S, \mathcal{T}$  do not make any sense in that limit. We thus need a further rewriting for which we need some more model-dependent input and to which we turn next.

## V. EXPANSION DETAILS

### A. McLennan formula: the first order

The first order in the expansion of the stationary distribution (13) has been known for a long time [30] and has been reconsidered more recently in [12, 26]. In fact, the idea of obtaining the McLennan-formula via the embedding described under Section III originates from [12]. The way how to deal with the limiting behavior  $T \uparrow \infty, \epsilon \downarrow 0$  was treated in [26]. We briefly repeat this and we concentrate on the Markov processes of Section II.

The main point is that

$$\langle S(\omega) \rangle_x^o = \epsilon \beta \int_0^T dt \langle w(x_t) \rangle_x^o \quad (18)$$

where, for jump processes  $w(x) = \sum_y k_o(x, y) f(x, y)$ , and for overdamped diffusions  $w = \chi \nabla \cdot f / \beta - \chi f \cdot \nabla U$ , see [26]. The expression (18) allows the limit  $T \uparrow \infty$  uniformly in  $\epsilon$  once we assume the equilibrium process to be irreducible and exponentially ergodic.

Plugging (18) into (16) thus gives the linear order expression

$$\rho(x) / \rho_o(x) = 1 - \epsilon \beta h(x) + O(\epsilon^2) \quad (19)$$

with

$$h(x) = \int_0^\infty dt \langle w(x_t) \rangle_x^o$$

in which the integral is exponentially convergent. The relation with local equilibrium distributions is also discussed in [26].

### B. Second order

We look at the  $m = 2$  term in (16). The main object to consider for jump processes is

$$\int_0^T ds \left\langle \sum_t^T f(x_{t-}, x_t) \sum_y k_o(x_s, y) f(x_s, y) \right\rangle_x^o$$

where the sum is over the jump times. On the other hand, for overdamped diffusions we can introduce  $\sigma(x_t, dx_t) = \epsilon \beta dx_t \circ f(x_t)$  and  $\mathcal{T}_1(x) = -\beta \epsilon f(x) \cdot \chi \nabla U + \epsilon \chi \nabla \cdot f$ , so that we must

deal then with

$$h_2(x) \equiv \int_0^T ds \int_0^T \langle \sigma(x_t, dx_t) \mathcal{T}_1(x_s) \rangle_x^o \quad (20)$$

We can keep with this “diffusion-”notation also for the “jump-”case if we then take  $\sigma(x_t, dx_t) = \epsilon\beta d\mathcal{N}_t f(x_{t-}, x_t)$  where  $\mathcal{N}_t$  is the Poisson process counting jumps, and  $\mathcal{T}_1(x) = \epsilon\beta \sum_y k_o(x, y) f(x, y)$ . We give the argument for the jump-case.

Because we are dealing with a correlation function of two time-extensive quantities, we have to use twice the exponential convergence to the equilibrium expectation. We will use the bound  $|\langle f(x_u)g(x_0) \rangle^o - \langle f(x_u) \rangle^o \langle g(x_0) \rangle^o| \leq C_f C_g e^{-\alpha u}$ , for some positive  $\alpha$  and where  $C_f$  and  $C_g$  bound the functions  $f$ , respectively  $g$ .

To this end, we split the  $s$ -integral in (20) over  $[0, t] \cup [t, T]$ . For the first part,

$$\int_0^T \int_0^t ds \langle \sigma(x_t, dx_t) \mathcal{T}_1(x_s) \rangle_x^o = \int_0^T dt \int_0^t ds \langle \mathcal{T}_1(x_t) \mathcal{T}_1(x_s) \rangle_x^o$$

as follows from (11). For the second part we use

$$\int_0^T \int_t^T ds \langle \sigma(x_t, dx_t) \mathcal{T}_1(x_s) \rangle_x^o \int_0^T dt \int_0^t \langle \sigma(x_s, dx_s) \mathcal{T}_1(x_t) \rangle_x^o$$

so that always  $s < t$  in what follows. Therefore,

$$h_2(x) = \int_0^T dt \int_0^t \langle [\sigma(x_s, dx_s) + \mathcal{T}_1(x_s) ds] \mathcal{T}_1(x_t) \rangle_x^o \quad (21)$$

First of all, the integrand has an equilibrium expectation equal to zero, as one can see from time-reversal invariance of equilibrium:

$$\langle [\sigma(x_s, dx_s) + \mathcal{T}_1(x_s) ds] \mathcal{T}_1(x_t) \rangle^o = \langle [-\sigma(x_t, dx_t) + \mathcal{T}_1(x_t) dt] \mathcal{T}_1(x_s) \rangle^o$$

As  $t$  is always bigger than  $s$  in the integrand of (21), we can do the same trick as before by replacing  $\langle \sigma(x_t, dx_t) \mathcal{T}_1(x_s) \rangle^o$  by  $\langle \mathcal{T}_1(x_t) dt \mathcal{T}_1(x_s) \rangle^o$ . This shows us indeed that the equilibrium expectation of the integrand in (21) is zero. We can make that more explicit by substituting the expressions for the jump-case, to have

$$\begin{aligned} \langle [\sigma(x_s, dx_s) + \mathcal{T}_1(x_s) ds] \mathcal{T}_1(x_t) \rangle_x^o = \\ (\epsilon\beta)^2 \langle [d\mathcal{N}_s f(x_{s-}, x_s) + \zeta(x_s) ds] \langle \zeta(x_t) | x_s \rangle_x^o \rangle_x^o \end{aligned} \quad (22)$$

for  $\zeta(x) \equiv \sum_y k_o(x, y) f(x, y)$ . Therefore, with  $p_s(x, y)$  the transition probability for the detailed balance reference dynamics to find  $y$  at time  $s$  when starting from  $x$ ,

$$\begin{aligned} \langle [\sigma(x_s, dx_s) + \mathcal{T}_1(x_s) ds] \mathcal{T}_1(x_t) \rangle_x^o = \\ (\epsilon\beta)^2 ds \sum_{y,z} [p_s(x, z) k_o(z, y) f(z, y) + p_s(x, y) k_o(y, z) f(y, z)] \langle \zeta(x_{t-s}) \rangle_y^o \end{aligned} \quad (23)$$

which indeed vanishes upon replacing  $p_s(x, y) \rightarrow \rho_o(y)$  by the equilibrium distribution because  $\rho_o(z) k_o(z, y) f(z, y) + \rho_o(y) k_o(y, z) f(y, z) = 0$  by detailed balance.

Furthermore, the equilibrium expectation of  $\mathcal{T}_1$  is zero,  $\langle \mathcal{T}_1(x_t) \rangle^o = 0$ , so that

$$|\langle \zeta(x_{t-s}) \rangle_y^o| \leq C(y) e^{-\alpha(t-s)}$$

The rest of the argument is straightforward by using that  $p_s(x, y)$  is exponentially close to  $\rho_o(y)$  as a function of the time  $s$ :

$$|h_2(x)| \leq \int_0^T dt \int_0^t ds e^{-\alpha(t-s)} e^{-\alpha' s} C'(x)$$

which is clearly finite as  $T \uparrow \infty$ .

## VI. APPLICATION TO NONLINEAR RESPONSE

One of the very first applications of the McLennan formula (the first order as in (19)) is the derivation of linear response around equilibrium. It is indeed possible to derive the Green-Kubo relations from it; see e.g. Section IIIB in [26]. The analogue can be done also for higher order response as we now indicate.

We imagine the system in equilibrium up to time zero. At that time an external stimulus  $\epsilon f$  is added, as modeled in the set-up of Section II, driving the system away from equilibrium. We can then estimate, using (16) say to second order in  $\epsilon$ , for an observation  $Q$  at time  $T$ ,

$$\begin{aligned} \langle Q(x_T) \rangle_{\rho_0}^\epsilon &= \langle Q \rangle^o - \langle Q(\pi x_0) S(\omega) \rangle^o + \frac{1}{2} \langle Q(\pi x_0) S(\omega) \mathcal{T}_1(\omega) \rangle^o + o(\epsilon^2) \\ &= \langle Q \rangle^o + \langle Q(x_T) S(\omega) \rangle^o - \frac{1}{2} \langle Q(x_T) S(\omega) \mathcal{T}_1(\omega) \rangle^o + o(\epsilon^2) \end{aligned}$$

When the perturbation  $\epsilon f$  is the difference or the gradient of a potential  $\epsilon V$ , then the entropy flux  $S = \beta \epsilon [V(x_T) - V(x_0)]$ , assuming that  $V(x) = V(\pi x)$  (see formulae (2) and (4)). The expectation value then reads, up to second order,

$$\begin{aligned} \langle Q(x_T) \rangle_{\rho_0}^\epsilon &= \langle Q \rangle^o + \beta \epsilon \langle Q(x_T) [V(x_T) - V(x_0)] \rangle^o \\ &\quad - \frac{\beta^2 \epsilon^2}{2} \int_0^T ds \langle Q(x_T) [V(x_T) - V(x_0)] L_o V(x_s) \rangle^o \end{aligned} \quad (24)$$

The first order term is consistent with the (equilibrium) fluctuation-dissipation theorem. In the second order term in (24) we have used the expression for the backward generator  $L_o V(x) =$

$\sum_y k_o(x, y)[V(y) - V(x)]$  for jump processes,  $L_o V(x) = -\chi \nabla U \cdot \nabla V + \chi \Delta V / \beta$  for diffusion processes, to rewrite

$$\mathcal{T}_1(\omega) = \epsilon \beta \int_0^T ds L_o V(x_s)$$

when  $f(x, y) = V(y) - V(x)$  (in the jump-case for  $\psi_\epsilon(x, y) = 1$  in (3)) or when  $f(x) = \nabla V(x)$  (in the diffusion-case (2)). As a consequence,

$$\left. \frac{\partial^2}{\partial \epsilon^2} \langle Q(x_T) \rangle_{\rho_0}^\epsilon \right|_{\epsilon=0} = -\frac{\beta^2}{2} \int_0^T ds \langle [Q(x_T) - Q(\pi x_0)][V(x_T) - V(x_0)] L_o V(x_s) \rangle^o$$

Observe that the terms in the expansion (24) are not just expectations of observables at one fixed time; they are correlation functions. It is therefore natural to investigate also the perturbation expansion of correlation functions in the nonequilibrium process. That follows most generally from expanding the exponential and the action in (6)–(7). It is then straightforward to check e.g. that

$$\left. \frac{\partial^2}{\partial \epsilon^2} \langle Q(x_t) \rangle_{\rho_0}^\epsilon \right|_{\epsilon=0} = \beta \frac{\partial}{\partial \epsilon} \langle \langle [Q(x_t) - Q(\pi x_0)][V(x_t) - V(x_0)] \rangle_{\rho_0}^\epsilon \rangle_{\epsilon=0} \quad (25)$$

The treatment of nonlinear response can be started differently, for example from applying Ward-Takahashi type identities starting from the fluctuation symmetry in the distribution of the entropy flux, see e.g. Section 10 in [23] or more recently in [1, 22, 29, 31, 35]. The advantage of the present treatment however is that all the terms in the expansion are explicitly expressed as correlation functions in the full equilibrium reference process. On the other hand, we do not know how to extend our ideas to deterministic dissipative dynamics. There the steady state attractor has a lower dimension than the embedding space and the invariant density becomes singular. That aspect probably becomes less stringent for macroscopic systems, or, depending on the reduced description and the nature of the observables one can attempt a suitable projection on some smooth manifold. For a perturbation theory featuring a first order term producing the standard fluctuation-dissipation theorem and higher order terms giving rise to nonlinear response, see e.g. [3] and [32] for two different approaches. For further convergence of ideas, it would be very helpful to identify the notion of dynamical activity within the thermodynamic formalism of smooth dynamical systems. One natural guess would proceed via the escape rate formalism as in [6].

## VII. CONCLUSION

Expansions for nonequilibrium mesoscopic systems can be made in various ways. If one is interested in the stationary distribution away from detailed balance, as we are, we can try

Born-type or Dyson-type perturbation expansions starting from the Master equation. These can certainly be computationally useful, but that however is not the main point of this present paper. What was attempted here is to give an expansion in terms of few (basically two) path-dependent physical observables, the entropy flux and the dynamical activity. It is in the same spirit of many attempts dating already from the 1950-60's where a major issue was to understand whether the stationary nonequilibrium distribution could be described in terms of macroscopic parameters only, [18, 19, 30]. Here we deal with mesoscopic systems and we incorporate more recent ideas concerning the importance of the kinetics (noise and activity) for finding the relative probability of states, [17]. The utility of the expansion depends on the relevance of nonlinear response around equilibrium for stochastic Markovian dynamics, and whether natural phenomena obey the various conditions imposed.

The starting point of our analysis is a path-integral formulation for the distribution of histories in terms of a reference equilibrium distribution. The breaking of detailed balance introduces irreversible fluxes, making up the time anti-symmetric term in the action, and also creates excesses in dynamical activity, governing the time-symmetric fluctuations. Both fluxes and activities can be written down explicitly as function of the system's trajectory for Markov jump and diffusion processes, where the physical interpretation is supported by the condition of local detailed balance. The subsequent expansion then mimics a high-temperature expansion for Gibbs distributions but now on space-time, where the exponential function gets Taylor-expanded. Additional symmetries simplify the expansion and fast relaxation properties enable full control over the terms in the expansion, as explicitly demonstrated for the first few. It remains however an interesting mathematical challenge to rigorously prove the convergence of the complete perturbation series under an appropriate minimal set of conditions on the dynamics.

## VIII. ACKNOWLEDGEMENTS

M.C. acknowledges support from the Swiss National Science Foundation (SNSF). C.M. benefits from the Belgian Interuniversity Attraction Poles Programme P6/02. B.W. was junior research fellow at the FWO, Flanders until 30-09-2010.

## Appendix A: Continuum limit of dynamical activity

There is a large literature concerning the derivation of Langevin dynamics or the Fokker-Planck equation from the Master equation for jump processes. Most of the time one develops around some small parameter like the inverse volume or the inverse number of particles in what is known as the Kramers-Moyal expansion, [10]. We do not repeat that here, but since the dynamical activity is still relatively unknown we show how to obtain (2) from (5) in a continuum limit.

Consider a small mesh  $\delta > 0$  and a walker on a ring with detailed balance transition rates

$$k_o(x, x \pm \delta) = D \exp -\frac{\beta}{2}[U(x \pm \delta) - U(x)], \quad D > 0$$

and with driven rates

$$k(x, x \pm \delta) = k_o(x, x \pm \delta) \exp \frac{\beta \epsilon \delta}{2} f(x, x \pm \delta)$$

The excess in dynamical activity when in state  $x$  is then, following (5),

$$k(x, x + \delta) - k_o(x, x + \delta) + k(x, x - \delta) - k_o(x, x - \delta)$$

which we expand to order  $\delta^2$  as

$$\begin{aligned} & D \left(1 - \frac{\beta \delta}{2} U'(x)\right) \left[\frac{\beta \epsilon \delta}{2} f(x, x + \delta) + \frac{\beta^2 \epsilon^2 \delta^2}{8} f^2(x, x + \delta)\right] \\ + & D \left(1 + \frac{\beta \delta}{2} U'(x)\right) \left[\frac{\beta \epsilon \delta}{2} f(x, x - \delta) + \frac{\beta^2 \epsilon^2 \delta^2}{8} f^2(x, x - \delta)\right] \\ = & \frac{D \beta^2 \epsilon^2 \delta^2}{8} [f^2(x, x + \delta) + f^2(x, x - \delta)] + \frac{D \beta \epsilon \delta}{2} [f(x, x + \delta) - f(x - \delta, x)] \\ - & \frac{D \beta^2 \delta^2 \epsilon}{4} U'(x) [f(x, x + \delta) + f(x - \delta, x)] \end{aligned}$$

In other words, setting also  $D\beta = \chi$  and  $f(x, x) = f(x)$ ,

$$\begin{aligned} & \lim_{\delta \downarrow 0} \frac{2}{\delta^2} [k(x, x + \delta) - k_o(x, x + \delta) + k(x, x - \delta) - k_o(x, x - \delta)] = \\ = & \frac{\chi \beta}{2} \epsilon^2 f^2(x) + \chi \epsilon f'(x) - \chi \beta \epsilon f(x) U'(x) \end{aligned} \tag{A1}$$

in which we recognize the instantaneous excess in dynamical activity (2).

## Appendix B: Underdamped diffusion processes

We consider here a Langevin dynamics for a particle with mass  $m$ , position  $q_t$  and velocity  $v_t$ :

$$\begin{aligned} dq_t &= v_t dt \\ m dv_t &= [F(q_t) - m\gamma v_t] dt + \sqrt{2D} dB_t \end{aligned} \quad (\text{B1})$$

$\gamma$  is the friction coefficient and  $B_t$  is a standard Wiener process giving rise to Gaussian white noise. The (symmetric) matrix  $D$  governs the variance of that noise. The Einstein relation between  $\gamma$  and  $D$  brings in the inverse temperature:  $\gamma = \beta D$ . The corresponding Fokker-Planck equation for the evolution of distributions  $\mu_t$  is

$$\frac{d}{dt}\mu_t + \nabla \cdot J_{\mu_t} = 0 \quad (\text{B2})$$

for  $\nabla = (\nabla_q, \nabla_v)$  and for current  $J_{\mu} = (J_{\mu}^q, J_{\mu}^v)$  with

$$J_{\mu}^q = mv\mu, \quad J_{\mu}^v = F\mu - \gamma mv\mu - D\nabla_v\mu \quad (\text{B3})$$

to be understood with diagonal matrices  $\gamma$  and  $D$ . As before we write the force  $F = \epsilon f - \nabla_q U$ . The equilibrium case  $\epsilon = 0$  has the equilibrium distribution

$$\rho_o(q, v) \propto e^{-\beta[m\frac{v^2}{2} + U(x)]}$$

Here the action is

$$\begin{aligned} A(\omega) &= -\log \frac{d\mathcal{P}^{\epsilon}(\omega)}{d\mathcal{P}^0(\omega)} \\ &= \mathcal{T}(\omega) - S(\omega) \end{aligned}$$

with

$$\mathcal{T}(\omega) = \frac{\epsilon^2}{2} \int_0^T dt f \cdot D^{-1} f - \epsilon \int_0^T dt f \cdot D^{-1} \nabla U - m \epsilon \int_0^T dv_t \circ D^{-1} f \quad (\text{B4})$$

$$S(\omega) = \epsilon \beta \int_0^T dt v \cdot f \quad (\text{B5})$$

One checks that  $S$  equals the work done by the nonconservative force  $f$ , times  $\beta$ . The dynamical activity  $\mathcal{T}$  consists of several terms with less obvious physical meanings. Still, it consists of in principle measurable quantities: forces on the one hand and  $D$  on the other hand, which depends on the friction coefficient and the mass of the particle. This dynamical activity turns up in the linear response around nonequilibrium, see [2]. A natural continuation of Appendix A would be to understand the Smoluchowski limit of (B4). We don't do that here but the expansion of Section IV and the formula (16) remain unchanged.

## References

---

- [1] D. Andrieux and P. Gaspard: A fluctuation theorem for currents and non-linear response coefficients, *J. Stat. Mech: Theor. Exp.* (2007) P02006.
- [2] M. Baiesi, E. Boksenbojm, C. Maes and B. Wynants: Nonequilibrium Linear Response for Markov Dynamics, II: Inertial Dynamics, *J. Stat. Phys.* **139**, 492-505 (2010).
- [3] G. Boffetta, G. Lacorata, S. Musacchio and A. Vulpiani: Relaxation of finite perturbations: Beyond the fluctuation-response relation, *Chaos* **13**, 806–811 (2003).
- [4] G. Crooks: Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, *Phys. Rev. E* **60**, 2721-2726 (1999).
- [5] B. Derrida: Non-equilibrium steady states: fluctuations and large deviations of the density and of the current, *J. Stat. Mech.* P07023 (2007).
- [6] P. Gaspard: *Chaos, Scattering and Statistical Mechanics*, Cambridge Nonlinear Science Series **9**, Cambridge University Press, 1998.
- [7] S. Goldstein, R. Kuik, J.L. Lebowitz and C. Maes: From PCA to Equilibrium System and Back, *Commun. Math. Phys.* **125**, 71–79 (1989).
- [8] V. Jaksic, C.-A. Pillet and L. Rey-Bellet: Entropic Fluctuations in Statistical Mechanics I. Classical Dynamical Systems, [arXiv:1009.3248v2 \[math-ph\]](https://arxiv.org/abs/1009.3248v2).
- [9] W. P. Johnson: The Curious History of Faa di Bruno’s Formula, *Amer. Math. Monthly* **109**, 217–234 (2002).
- [10] N.G. van Kampen: *Stochastic processes in physics and chemistry*, North Holland 2007.
- [11] S. Katz, J.L. Lebowitz, and H. Spohn: Stationary nonequilibrium states for stochastic lattice gas models of ionic superconductors, *J. Stat. Phys.* **34**, 497-538 (1984).
- [12] T. S. Komatsu and N. Nakagawa: An expression for stationary distribution in nonequilibrium steady states, *Phys. Rev. Lett.* **100**, 030601 (2008).
- [13] T. S. Komatsu, N. Nakagawa, S. Sasa, and H. Tasaki: Representation of nonequilibrium steady states in large mechanical systems, *J. Stat. Phys.* **134**, 401–423 (2009).
- [14] T. S. Komatsu, N. Nakagawa, S. Sasa, and H. Tasaki: Steady State Thermodynamics for Heat Conduction – Microscopic Derivation, *Phys. Rev. Lett.* **100**, 230602 (2008).

- [15] R. Landauer: Inadequacy of entropy and entropy derivatives in characterizing the steady state, *Phys.Rev. A* **12**, 636–638 (1975).
- [16] R. Landauer: Stability and entropy production in electrical circuits, *J. Stat. Phys.* **13**, 1-16 (1975).
- [17] R. Landauer: Motion Out of Noisy States, *J. Stat. Phys.* **53**, 233–248 (1988).
- [18] J. L. Lebowitz and P. G. Bergmann: Irreversible Gibbsian Ensembles, *Annals of Physics* **1**, 1 (1957).
- [19] J. L. Lebowitz: Stationary Nonequilibrium Gibbsian Ensembles, *Phys. Rev.* **114**, 1192–1202 (1959).
- [20] J. L. Lebowitz, C. Maes and E.R. Speer: Statistical Mechanics of Probabilistic Cellular Automata, *J. Stat. Phys.* **59**, 117–170 (1990).
- [21] V. Lecomte, C. Appert-Rolland, and F. van Wijland: Thermodynamic formalism for systems with Markov dynamics, *J. Stat. Phys.* **127**, 51–106 (2007).
- [22] E. Lippiello, F. Corberi, A. Sarracino and M. Zannetti: Nonlinear response and fluctuation dissipation relations, *Phys. Rev. E* **78**, 041120 (2008) .
- [23] C. Maes: On the origin and the use of fluctuation relations for the entropy, *Séminaire Poincaré* **2**, 29–62 (2003).
- [24] C. Maes and M. H. van Wieren: Time-symmetric fluctuations in nonequilibrium systems, *Phys. Rev. Lett.* **96**, 240601 (2006).
- [25] C. Maes and K. Netočný: Time reversal and entropy, *J. Stat. Phys.* **110**, 269–310 (2003).
- [26] C. Maes and K. Netočný: Rigorous meaning of McLennan ensembles, *J. Math. Phys.* **51**, 015219 (2010).
- [27] C. Maes, K. Netočný, and B. Wynants: On and beyond entropy production: the case of Markov jump processes, *Markov Processes Relat. Fields* **14**, 445–464 (2008).
- [28] C. Maes, F. Redig and A. Van Moffaert: On the definition of entropy production, via examples, *J. Math. Phys.* **41**, 1528–1554 (2000).
- [29] U. M. B. Marconi, A. Puglisi, L. Rondoni and A. Vulpiani: Fluctuation-Dissipation: Response Theory in Statistical Physics, *Phys. Rep.* **461**, 111 (2008).
- [30] J. A. McLennan Jr.: Statistical mechanics of the steady state, *Phys. Rev.*, **115**, 1405–1409 (1959).
- [31] K. Mallick, M. Moshe and H. Orland: A field-theoretic approach to nonequilibrium work identities, [arXiv:1009.4800](https://arxiv.org/abs/1009.4800).
- [32] D. Ruelle: A review of linear response theory for general differentiable dynamical systems, *Nonlinearity* **22**, 855–870 (2009).
- [33] Ken Sekimoto: *Stochastic Energetics*, Lecture Notes in Physics **799**, Springer (2010).

- [34] H. Spohn: *Large Scale Dynamics of Interacting Particles*, Springer Verlag, Heidelberg, 1991.
- [35] D. Villamaina, A. Baldassarri, A. Puglisi and A. Vulpiani: The fluctuation-dissipation relation: how does one compare correlation functions and responses?, *J. Stat. Mech.: Theor. Exp.* P07024 (2009).